Community for Data Integration (CDI) FY19 Request for Proposals Process

Leslie Hsu, CDI Coordinator

October 24, 2018

1-2 pm ET

Topics

- Overview of CDI
- Background and overview of RFP
- New in FY19
- Examples of successful projects
- Overview of Submission and Voting
- Key Dates
- Questions

Why we have the CDI RFP

The purpose of the CDI RFP is to build USGS capabilities in data integration and management.

CDI projects.... are short-term | leverage existing resources | demonstrate scalable solutions | improve access to data and tools | develop best practices | share lessons | involve interdisciplinary teams | can be modified for reuse | communicate their value to the wider USGS.

The CDI RFP increases communication across boundaries and creates opportunities to work with people outside of your normal program.

The proposals process involves .... commenting and voting from your peers | new collaborations | communicating without disciplinary jargon | community-building.

CDI - Overview

- Chartered as Council on Data Integration in 2009 to:
  - Lead development of Data Integration Strategy
  - Recommend data integration guidelines
  - Promote USGS-wide data integration
- Evolved into an open Community of Practice

Executive Sponsors:
Kevin T. Gallagher (left) and Tim Quinn (right)

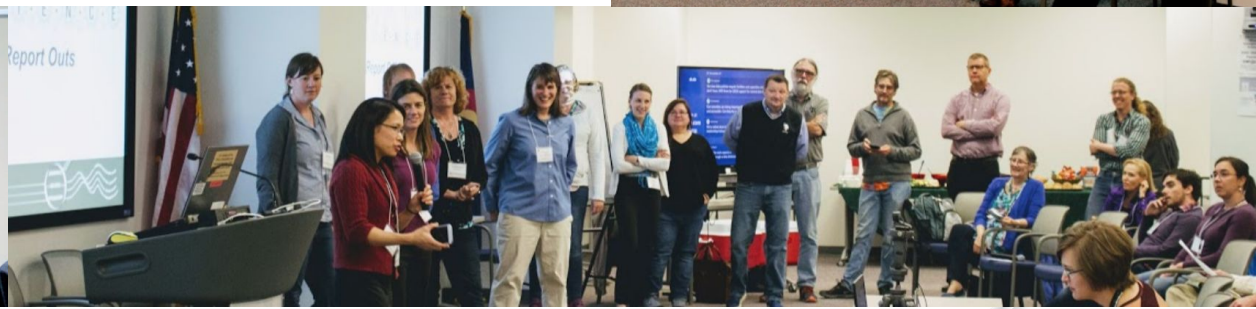
Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly – Etienne Wenger, 2011

The CDI Purpose

- Advance understanding of Earth systems through enhanced use of data and information
- Provide a forum for sharing ideas and learning new skills
- Increase the visibility of work with data and information

CDI Today

- **Monthly Forum**
  - 2nd Wednesday at 11 ET
- **Collaboration Areas**
- **Annual Workshop/Webinar**
- **RFP Process**



FY19 CDI RFP

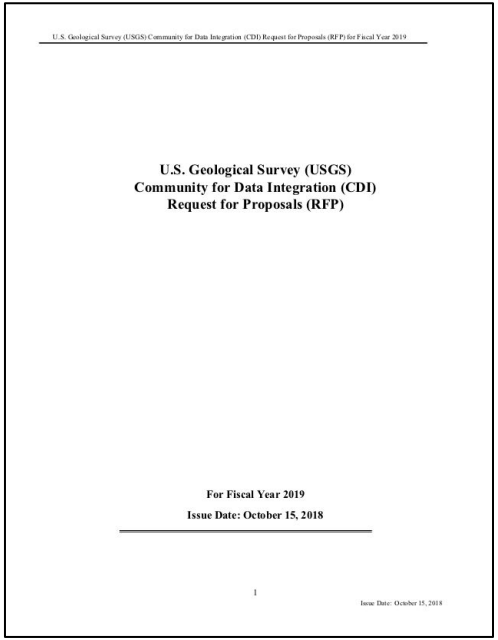
https://my.usgs.gov/confluence/display/cdi/2019+Proposals

2019 Proposals

Created by Hsu, Leslie, last modified on Oct 15, 2018

This page describes the Fiscal Year 2019 Community for Data Integration (CDI) Request for Proposals (RFP). Please watch this page for updates throughout the proposals process.

- Guidance Document
- Estimated Schedule
- Collaboration Forum
- Why do we have the Community for Data Integration Request for Proposals?
- How does the proposals process work?
- What kind of projects are successful?


U.S. Geological Survey (USGS) Community for Data Integration (CDI) Request for Proposals (RFP) for Fiscal Year 2019

U.S. Geological Survey (USGS)
Community for Data Integration (CDI)
Request for Proposals (RFP)

For Fiscal Year 2019
Issue Date: October 15, 2018

1

Issue Date: October 15, 2018

2 Phase RFP

Phase 1: Statement of Interest

- 1.5 page document, due November 16, 2018
- Relate to Science Support Framework, Guiding Principles, and Evaluation Criteria
- Lightning presentations to introduce ideas to the community (November 28, 2018)
- Commenting and voting by CDI Community Members
- Top ranked SOIs invited to submit full proposal

Phase 2: Invited Full Proposal

- Proposal narrative - 7 page maximum, not including plain language summary, detailed budget, data management planning form, letters of support, CVs
- Relate to Science Support Framework, Guiding Principles, and Evaluation Criteria
- Evaluated by formal Review Panel
- Recommendations presented to CDI Executive Sponsors for selection and award

Details of the FY19 RFP

- Eligibility: USGS employee must be PI, but collaboration is encouraged
- Projects complete in 3-6 months
- Mid Project and End of Project Reporting
- Required to include travel $ to the 2019 CDI Workshop in Boulder, CO (June 4-7, 2019).
- All results should be publicly available and follow USGS Fundamental Science Practices
- Topical themes: FAIR data, reuse of CDI products, building national datasets, and biosurveillance
- Proposals on any topic still accepted

FY19 RFP Themes

This year, the CDI executive sponsors are encouraging proposals that produce building blocks for an Integrated Predictive Science Capacity with the following themes.

- Producing FAIR (Findable, Accessible, Interoperable, and Reusable) data and tools for Integrated Predictive Science Capacity (see the Enabling FAIR Data site)
- Reusing or repurposing modular tools such as those that were developed by previous CDI projects, including the CDI Risk Map,
- Building authoritative national datasets for hazards or assets (integrating data and assessing quality),
- Tools and methods for biosurveillance of emerging invasive species and health threats.

Funding

- Max request = $50K
- CDI funds can be used for salary, travel, other expenses
- CDI funds CANNOT be used for field work or Core Science System Mission Area federal employees salary
- Proposal must have 30% matching funds from salary, travel, other grants, etc.

New in FY19

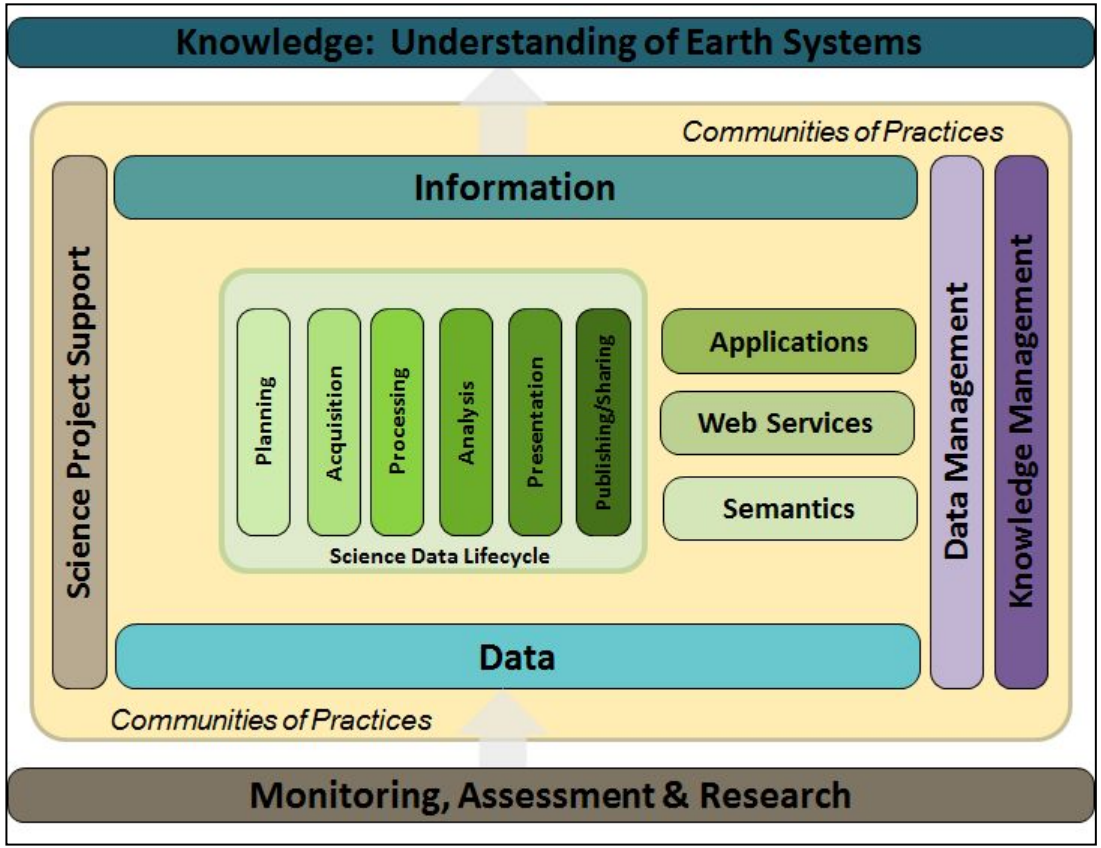
- **FY19 Themes** - FAIR data, re-using previous CDI products, building national datasets, biosurveillance
- **Collaboration Forum** - Method for finding collaborators - technical or disciplinary
- **Proposal Submission -** Uses Google Forms
- **Commenting and Voting -** Using the CDI wiki and the SimplyVoting app, not IdeaLab (a.k.a. UserVoice)
- **Lightning SOI Presentation Session**, November 28, 3pm ET - *one-slide template, a single image, and option to pre-record*
- **Examples of previous successful SOIs and Proposals** (link)

How to frame your idea

- **Science Support Framework (SSF)**
  - Applicants select top 3 SSF elements.
- **Guiding Principles**
- **Evaluation Criteria**

CDI Science Support Framework

The foundation for sharing science



Guiding Principles

- Focus on targeted efforts that yield near-term benefits to Earth and biological science
- Leverage existing capabilities and data
- Implement and demonstrate innovative solutions (e.g. methodologies, tools, or integration concepts) that could be used or replicated by others at scales from project to enterprise
- Preserve, expose, and improve access to Earth and biological science data, models, and other outputs
- Develop, organize, and share knowledge and best practices in data integration

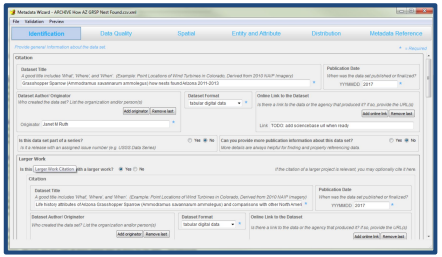
Evaluation Criteria

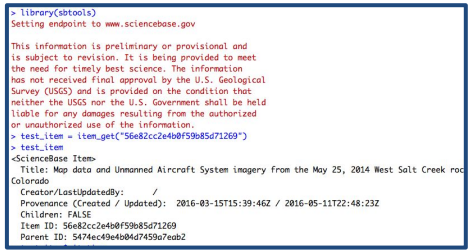
Scope (25%)	• Is there a demonstrated need for the effort/activity? • How much does the proposal contribute to the guiding principles and element(s) of the CDI Science Support Framework? • Is there potential impact beyond a single Program, Center, Mission Area, or Region? • What is the anticipated return on investment (e.g. cost savings, code utilization, publications, operational efficiencies, etc.)?
Technical Approach (25%)	• Is the technical approach applied to the problem reasonable? • Is the approach innovative? • Does it employ a proven, reliable technique that is appropriate to the problem?
Project Experience and Collaboration (25%)	• Does the project team have the appropriate experience, special qualifications, and skills for successful completion of the project.? • Have inter-disciplinary or cross-mission area/region collaboration and partnerships been pursued where appropriate?

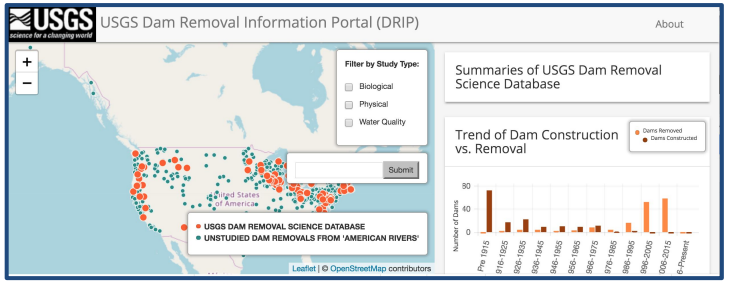
Evaluation Criteria (2)

Sustainability, Outreach, and Communication (15%)	• Does the proposal describe the intended sustainability of the project deliverables (products, tools, services, metadata) for long-term access, reusability, and potential for integration? • Will all products resulting from the project be freely shared and made available, without charge or restriction, to the CDI, the broader USGS community, and beyond as appropriate?
Budget Justification (5%)	• Is the budget at or below $50,000 and does it meet the minimum 30% in-kind match? • Are salaries and contractor costs, travel, and equipment/ publication costs justified and appropriate to project needs? • If the proposal lists contractor support, does the proposal describe how the contract work will be managed and documented?
Timeline (5%)	• Are the project phases and milestones described in the technical approach? • Is the proposed workload feasible given the project duration.

Example Past CDI Projects

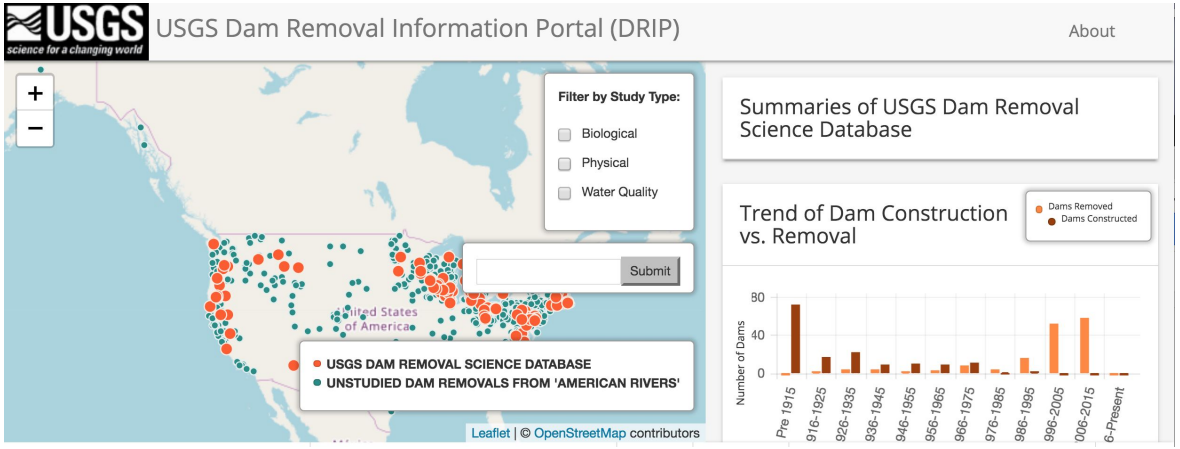






DRIP: Dam Removal Information Portal

- Funded in 2015
- OFR describing methods inspired other CDI projects to take similar steps to publish and expose datasets using existing infrastructure.
- Continues with related projects funded from other sources.



sbtools: an R package for ScienceBase

- Funded in 2015
- Tool to more easily access ScienceBase through R (command line)

- A command-line tool easy enough for the CDI coordinator to use!
- Increases accessibility to ScienceBase data, information, and services.

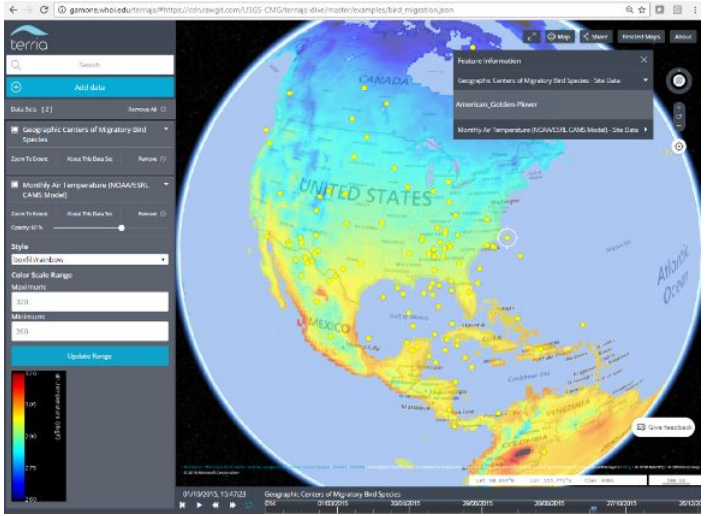
```
> library(sbtools)
Setting endpoint to www.sciencebase.gov

This information is preliminary or provisional and
is subject to revision. It is being provided to meet
the need for timely best science. The information
has not received final approval by the U.S. Geological
Survey (USGS) and is provided on the condition that
neither the USGS nor the U.S. Government shall be held
liable for any damages resulting from the authorized
or unauthorized use of the information.
> test_item = item_get("56e82cc2e4b0f59b85d71269")
> test_item
<ScienceBase Item>
  Title: Map data and Unmanned Aircraft System imagery from the May 25, 2014 West Salt Creek roc
Colorado
  Creator/LastUpdatedBy:      /
  Provenance (Created / Updated):  2016-03-15T15:39:46Z / 2016-05-11T22:48:23Z
  Children: FALSE
  Item ID: 56e82cc2e4b0f59b85d71269
  Parent ID: 5474ec49e4b04d7459a7eab2
```

terriaJS - Evaluating a new open-source, standards-based framework for web portal development

- Funded in 2016
- Assessed a framework for creating web portals by non-developers with the current USGS landscape of available tools.

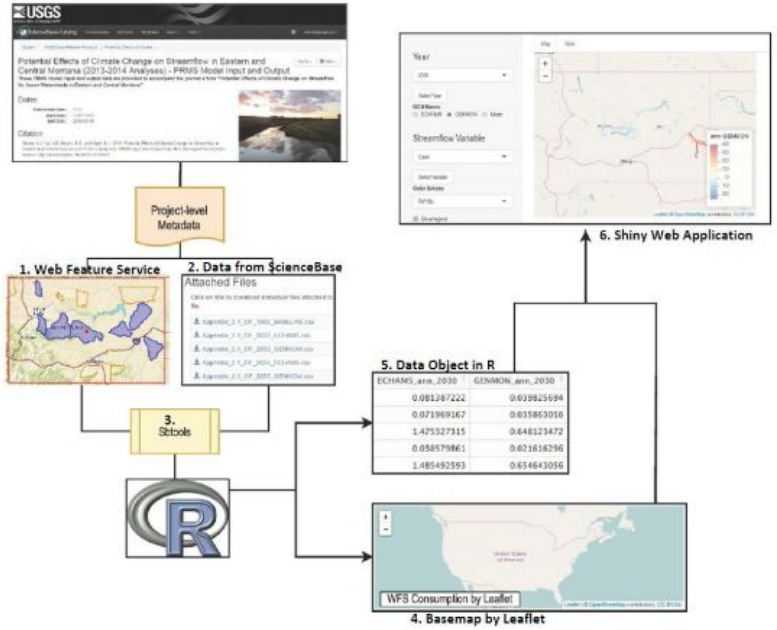
- Framework later used by multiple USGS groups including the CDI Risk Map Project.



Recommended practices for publishing digital data through Sciencebase

- Funded in 2016
- Documented processes for USGS scientists to organize and share data using ScienceBase, and to provide an example interactive mapping application to display those data

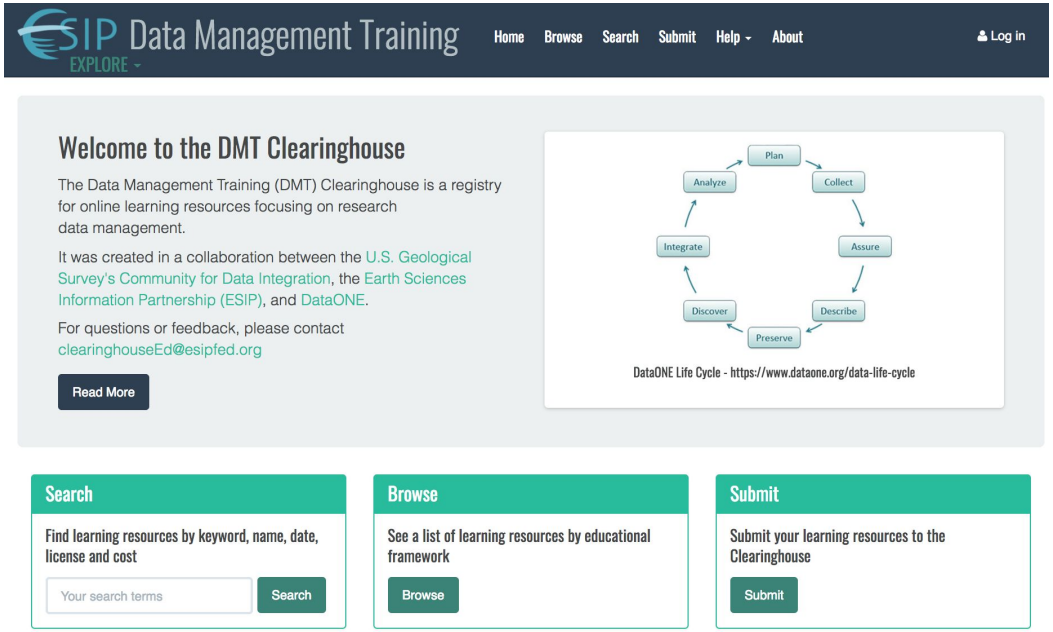
- Topic in high demand: webinar outside of CDI attracted over 100 participants.



Data Management Training Clearinghouse

- Funded in 2016
- Tool to discover and evaluate data management training modules.

- Now being used as a platform for the American Geophysical Union's FAIR Data project to organize training resources for scientists.



More information on past projects

Browse past projects to see what has been funded by CDI.

- ScienceBase CDI Community (by year)
- Annual Reports have project summaries*

*To access *all* of the materials, email cdi@usgs.gov to become a CDI member.

Proposal Concepts *not* CDI focused

- Supporting the collection of new data or field research
- Monitoring, assessment, or dataset creation projects. Although CDI may fund the creation of some broadly-usable (“foundational”) data content, this is normally considered out of scope
- Projects that would normally be funded by individual Program Areas or other proposal processes such as the John Wesley Powell Center for Analysis and Synthesis, Center of Excellence for Geographic Information Science (CEGIS), Innovation Center for Earth Sciences (ICES), and Office of Organizational and Employee Development (OED). (Explanation of different funding opportunities.)

New RFP Submission Process

- Submission by Google Form on the FY19 Proposals wiki page
- Prepare the information listed in Appendix A of the Guidance Document
- Use the template .docx, submit as a PDF
- Use the SOI checklist on the last page of the Guidance Document
- You can save your progress by clicking to the end and Submitting. You can update your information until the deadline on **November 16, 5pm ET**.

Tips and Observations

- Communicate potential for partnerships and reusability – Use the Collaboration Forum
- Visit the FAQ page
- Speak to a broad USGS audience - you are trying to gain support for your idea and must communicate its value to the voters
- Submission form tips
  - Save/update often, compose text on external text editor
- Budget – double check that you have followed all guidelines
  - Travel to the 2019 CDI Workshop, June 4-7 in Boulder, CO.
- Practice and time your 1-minute lightning talk - sometimes less is more. How will your project benefit USGS?

SOI Commenting and Voting Process

- Commenting will open the Monday after statements are due
- Lightning Presentations - Wednesday, **November 28**, 3pm ET
- Community voting through our CDI wiki
  - Must be a CDI Member by **Nov 29** (email to join: cdi@usgs.gov)
  - Consider all SOIs before voting
  - Community can comment on SOIs – important for PIs to respond
  - Rank top proposals
- Voting closes – **December 14**, 11:59pm ET
- Full proposals selected from all SOIs submitted. Based on variety of factors:
  - # of votes
  - Topical focus, USGS emerging priorities

Key Dates in FY19

November 16 @ 5 pm ET	**Statements of Interest (SOI) Due**
November 28 @ 3pm ET	SOI Lightning Presentation Session
November 30 - December 14	Voting Period
Early January 2019	Invite Full Proposal Submissions
Mid-February 2019	Full Proposals Due
March 2019	Awarded Projects Announced
September 30, 2019	Awarded Funds Spent

Questions?

Email: cdi@usgs.gov

Website: http://www.usgs.gov/cdi

RFP Announcement:

https://my.usgs.gov/confluence/display/cdi/2019+Proposals

Extra slides

Code Repositories

- Examples: Previous repos on GitHub or Bitbucket
- Gitlab recently available to USGS (e.g., code.usgs.gov)
- PIs will be given options to create repositories
- Repositories should be public after the completion of the project.

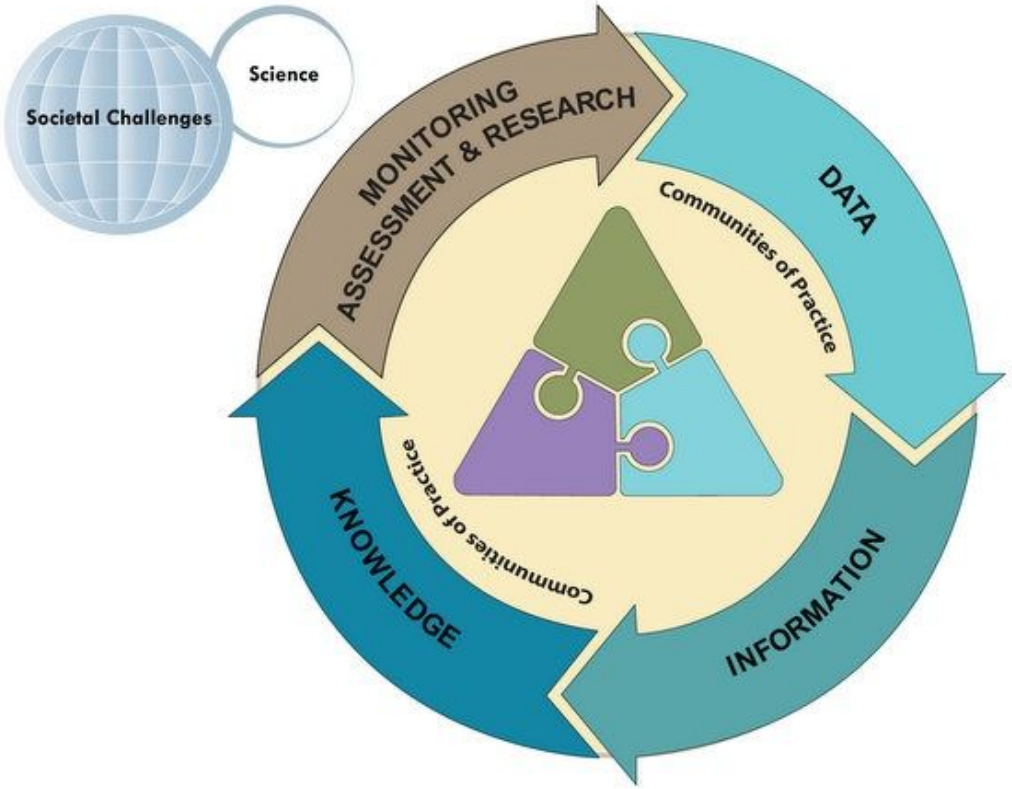
Big Picture

Communities of Practice

Computational Tools & Services

Management, Policy & Standards

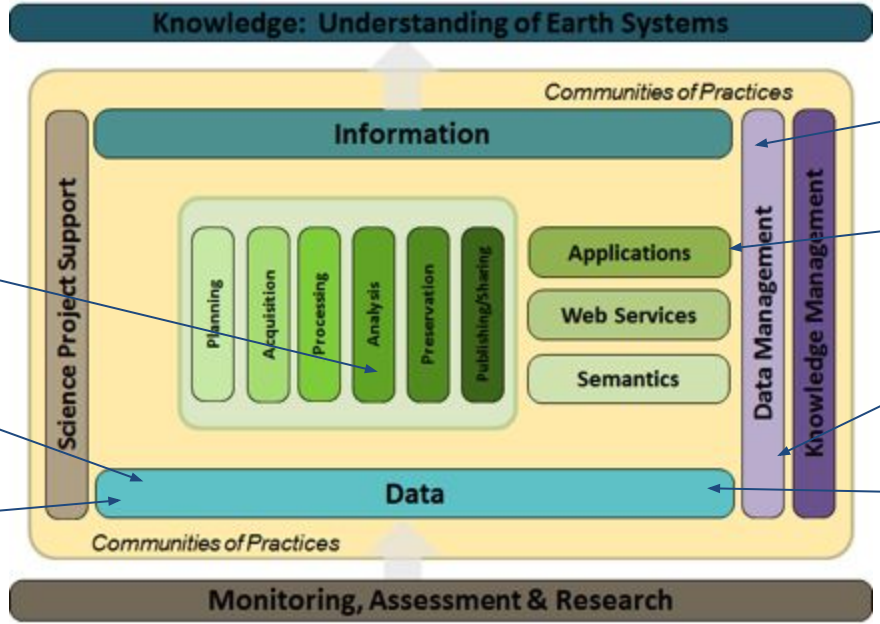
Data & Information Assets



Upper Level Categories

 Data & Information Assets	**Data** Persistent archives, data registries, catalogs, data, metadata, derived information products, knowledge bases, vocabularies/ontologies.
 Computational Tools & Services	**Technology** Applications, web services, data discovery tools, models, semantic services & tools, infrastructure, data brokers and visualization tools.
 Management, Policy & Standards	**Business** Data stewardship, implementation of the Data Management Life Cycle, knowledge management, data standards, governance and policy.
 Communities of Practice	**People** Scientists, CDI as a whole, CDI Working Groups, external partners, and the human network of scientific domain collaborators.

CDI Science Support Framework - FY18 Projects



Mapping with deep neural networks

Knowledge Extraction Algorithms (KEA)

Long-term Drought Risk

Transition to ISO metadata

Forecasting invasive species

Ice Jam Hazard Mobile-Enabled Website

National Alert Risk Mapper for NAS